

# A NEW MODE REDUCTION STRATEGY FOR THE GENERALIZED KURAMOTO-SIVASHINSKY EQUATION

M. SCHMUCK <sup>\*</sup>, M. PRADAS <sup>†</sup>, G. A. PAVLIOTIS <sup>‡</sup>, AND S. KALLIADASIS <sup>§</sup>

**Abstract.** Consider the generalized Kuramoto-Sivashinsky (gKS) equation. It is a model prototype for a wide variety of physical systems, from flame-front propagation, and more general front propagation in reaction-diffusion systems, to interface motion of viscous film flows. Our aim is to develop a systematic and rigorous low-dimensional representation of the gKS equation. For this purpose, we approximate it by a renormalization group equation which is qualitatively characterized by rigorous error bounds. This formulation allows for a new stochastic mode reduction guaranteeing optimality in the sense of maximal information entropy. Herewith, noise is systematically added to the reduced gKS equation and gives a rigorous and analytical explanation for its origin.

These new results would allow to reliably perform low-dimensional numerical computations by accounting for the neglected degrees of freedom in a systematic way. Moreover, the presented reduction strategy might also be useful in other applications where classical mode reduction approaches fail.

**Key words.** Generalized Kuramoto-Sivashinsky equation, renormalization group method, stochastic mode reduction

**Subject classifications.** 35Q53, 37F25, 81T15, 81T17

**1. Introduction** We consider abstract evolution equations of the form,

$$\begin{aligned} \frac{\partial u}{\partial t} + B(u, u) + Au &= 0, \\ u(x, 0) &= u_0(x), \end{aligned} \quad (1.1)$$

where  $A$  denotes a general linear operator and  $B$  represents a nonlinear term of Burgers' type. Well known equations in this class include, e.g. the viscous Burgers equation, the Korteweg-de Vries equation, and the Benney-Lin equation. We start by performing a formal renormalization group (RG) approach for the general form in (1.1). We subsequently focus on a rigorous low-dimensional reduction of the generalized Kuramoto-Sivashinsky (gKS) equation:

$$\begin{aligned} \partial_t u + \lambda u u_x + \kappa u_{xx} + \delta u_{xxx} + \nu u_{xxxx} &= 0, & \text{in } \mathcal{P}_\alpha \times ]0, T[, \\ u(x, 0) &= g(x), & \text{in } \mathcal{P}_\alpha, \end{aligned} \quad (1.2)$$

where  $\mathcal{P}_\alpha := ]-\alpha\pi, \alpha\pi[$  is a periodic domain with  $\alpha := \frac{L}{2\pi}$  for an arbitrary period  $L > 0$  while the solution  $u(x, t) : \mathcal{P}_\alpha \times ]0, T[ \rightarrow \mathbb{R}$  of (1.2) describes for example the height of a one-dimensional surface above a substrate point  $x$  at time  $t$ . We also take  $g(x) \in H^q(\mathcal{P}_\alpha)$  for  $q \geq 4$ , a periodic initial condition, i.e.,

$$g(x + L) = g(x).$$

<sup>\*</sup>Departments of Chemical Engineering and Mathematics, Imperial College, London SW7 2AZ, UK (m.schmuck@imperial.ac.uk),

<sup>†</sup>Department of Chemical Engineering, Imperial College, London SW7 2AZ, UK (m.pradas@imperial.ac.uk)

<sup>‡</sup>[INRIA MicMac Project Team, Ecole des Ponts ParisTech, 6 et 8 Avenue Blaise Pascal, Cité Descartes Champs sur Marne, 77455 Marne la Vallée Cedex 2, France], on leave from [Department of Mathematics, Imperial College, London SW7 2AZ, UK (g.pavliotis@imperial.ac.uk)]

<sup>§</sup>Department of Chemical Engineering, Imperial College, London SW7 2AZ, UK (s.kalliadasis@imperial.ac.uk).

The gKS equation is of the type (1.1) with,

$$B(u) := B(u, u) := \lambda u u_x, \quad \text{and} \quad A := (\kappa \partial_x^2 + \delta \partial_x^3 + \nu \partial_x^4). \quad (1.3)$$

A rigorous dimension reduction of the gKS equation is of special interest because it is not a Hamiltonian system and does not have an intrinsic invariant measure. This makes direct application of stochastic mode reduction strategies difficult, see [33] for instance.

It is noteworthy that the gKS equation retains the fundamental elements of any nonlinear process that involves wave evolution: the simplest possible nonlinearity  $uu_x$ , instability and energy production  $u_{xx}$ , stability and energy dissipation  $u_{xxxx}$  and dispersion  $u_{xxx}$ . We notice that the nonlinearity arises effectively from the nonlinear correction to the phase speed, a nonlinear kinematic effect that captures how larger waves move faster than smaller ones. In the context of thin-film flows [6, 36, 37], the terms  $uu_x$ ,  $u_{xx}$ ,  $u_{xxx}$  and  $u_{xxxx}$  are due to the interfacial kinematics associated with mean flow, inertia, viscosity and surface tension, respectively, with the corresponding parameters,  $\lambda$ ,  $\kappa$ ,  $\delta$  and  $\nu$  all positive and measuring the relative importance of these effects. The “strength of the nonlinearity”,  $\lambda$ , in particular, is associated with the scaling for the velocity (and hence time). In addition,  $\int_{\mathcal{P}_a} u dx = u_0$ , a measure of the volume of the liquid, a conservation property for systems whose spatial average does not drift. A simplified form of (1.2) is obtained by appropriately rescaling  $u$ ,  $x$  and  $t$  which is equivalent to setting  $\lambda = \kappa = \nu = 1$  and keeping the same notation for dimensionless quantities [6, 36, 37].

As with many nonlinear time-dependent problems in science and engineering, equations of the form (1.1) are too complex to be fully resolved and the influence of neglected degrees of freedom is not clear a priori. This problem exists independently of spatial dimensions for (1.1) and hence for the gKS equation also. The reliable resolution of high dimensional problems is a well-known issue in computational science where one can numerically only deal with a finite number of degrees of freedom.

Hence, there is a strong need for dimensionally reduced formulations, which in turn would allow for studies of long time behavior of physical systems. Modeling of the ocean-atmosphere, which mainly generates our weather, is one important example: One has a characteristic timescale of several years for the ocean in contrast to a couple of days governing atmospheric structures such as cyclones. As a consequence, a characteristic feature of many physical systems is the presence of fast and slow degrees of freedom. The relevant information of a system’s long time behavior is often primarily contained in the slow modes.

This immediately leads to the general, often ad hoc, approximation of decomposing the problem of interest into fast  $w$  and slow  $v$  modes. For equation (1.1) such a purely formal splitting, i.e.,  $u \approx u^\epsilon = v^\epsilon + w^\epsilon$ , reads in standard notation applied in the literature as,

$$\begin{aligned} \frac{\partial}{\partial t} v^\epsilon &= f(v^\epsilon, w^\epsilon), \\ \frac{\partial}{\partial t} w^\epsilon &= \frac{1}{\epsilon} g(v^\epsilon, w^\epsilon), \end{aligned} \quad (1.4)$$

where the small parameter  $0 < \epsilon \ll 1$  mediates the timescale separation. Mode reduction strategies, such as “adiabatic elimination” [38], invariant manifolds [7], and optimal prediction [3], are tools to eliminate the fast modes and derive “appropriate”

equations for the slow modes only. We remark that especially for systems with spatio-temporal chaos (like the gKS equation [4, 36]) such a reduction needs to be carefully performed in order to not lose the relevant dynamical characteristics of the full system. The strategy of defining an invariant manifold is almost classical by now. For example, in [8] the existence of an inertial manifold for the KS equation (obtained from (1.2) with  $\delta = 0$ ) is shown. An inertial manifold is a finite-dimensional, exponentially attracting, positively invariant Lipschitz manifold. The principle idea is to determine a map  $\Phi: \mathbb{V} \rightarrow \mathbb{W}$  such that we can rewrite equation (1.1) in the low-dimensional form,

$$\partial_t v + \mathbf{P}\mathbf{B}(v + \Phi(v), v + \Phi(v)) + \mathbf{A}\mathbf{P}v = 0, \quad (1.5)$$

where  $\mathbf{P}: \mathbb{H} \rightarrow \mathbb{V}$  and  $\mathbf{Q} := (\mathbf{I} - \mathbf{P}): \mathbb{H} \rightarrow \mathbb{W}$  are projections onto the orthogonal subspaces  $\mathbb{V}$  and  $\mathbb{W}$  such that  $\mathbb{H} = \mathbb{V} \oplus \mathbb{W}$ . A strategy to determine  $\Phi$  in general Galerkin spaces is, for example, suggested in [7] for the KS equation. The RG approach performed here can also be understood as a formal and feasible procedure to derive an asymptotic invariant manifold, see (1.8) and (1.9).

**Open questions and answers to the classical separation (1.4):** (i) *Is the splitting (1.4) and the approximation of  $u$  by  $u^\epsilon$  valid and in which sense?* This question is often not answered in the literature where from the outset a separation (1.4) is assumed, see [1, 4] for instance. These studies heuristically motivate a timescale mediation (or separation into slow and fast scales) of the form (1.4). The present work aims to provide a rigorous foundation in Theorem 3.6 by the following estimate,

$$\|u - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)}^2(T) \leq C\epsilon^2 + \exp(CT) \left( \epsilon^{1/4} + \epsilon \right). \quad (1.6)$$

If we suppose that  $u$  and  $u^\epsilon$  satisfy a Gevrey regularity characterized by a parameter  $\sigma > 0$ , then we can improve (1.6) in the following way,

$$\|u - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)}^2(T) \leq C\epsilon^2 + \exp(CT) \left( \epsilon^{1/4} \exp\left(-\frac{\sigma}{\epsilon^{1/4}}\right) + \epsilon \right). \quad (1.7)$$

It should be pointed out that these estimates also account for the reduction to the slow degrees of freedom  $v^\epsilon$  too and not only for the error between  $u$  and  $u^\epsilon$ .

(ii) *How can we account for the fast degrees of freedom  $w^\epsilon$  in an equation for the slow modes  $v^\epsilon$  only?* For this purpose we apply an abstract RG approach extended to general multiscale problems, see [2, 24, 25]. The RG method was first introduced in quantum field theory as a tool to perform scale transformations. The method then became popular with Wilson's work on the Kondo problem [39]. It can formally provide the separation (1.4). This means, we first obtain an approximation for  $v^\epsilon$  of the form

$$\partial_t v^\epsilon + \mathbf{A}_v v^\epsilon + \mathbf{P}_N \mathbf{B}(v^\epsilon, v^\epsilon) = -\epsilon \mathbf{G}_\epsilon(U(t), v^\epsilon), \quad (1.8)$$

where  $\epsilon \mathbf{G}_\epsilon(U, v^\epsilon)$  is a perturbation “force” originating from the renormalization method and  $U = V + W = v^\epsilon + W$  is a solution of the RG equations

$$\begin{aligned} \partial_t V + \mathbf{A}_v V + \mathbf{P}_N \mathbf{B}(V, V) &= 0, \\ \partial_t W + \mathbf{Q}_N \mathbf{B}_1(V, W) &= 0, \end{aligned} \quad (1.9)$$

where  $V = \mathbf{P}_N U$  and  $W = (\mathbf{I} - \mathbf{P}_N)U =: \mathbf{Q}_N U$  are projections onto the slow and fast manifolds, respectively. Since we can analytically solve for  $W$ , we end up with an

equation for the slow variable  $v^\epsilon$  only. The above estimates (1.6) and (1.7) then make the reduction (1.8) rigorous. Moreover, equation (1.9)<sub>2</sub> can be interpreted as the map  $\Phi^\epsilon(v^\epsilon)$  onto the asymptotic invariant manifold.

It should also be pointed out that at this stage the RG approximation (1.8) alone is not satisfactory since the fast variable  $W$  contained in  $U$  is of infinite dimension and hence can not entirely be resolved numerically. We give an answer to this problem after the last question (iii) by the principle of maximum entropy. Moreover, question (ii) is of particular relevance here, since the fast modes prevent the existence of a canonically invariant measure. Such a measure is the key information for classical reduction methods such as Mori-Zwanzig and optimal prediction, see question iii).

(iii) *What kind of information do we need to carry over from the (infinite dimensional) fast degrees of freedom to the (finite dimensional) slow ones and how?* To this end, we derive a stochastic evolution equation for the resolved (slow) variable by properly including necessary information from the unresolved (fast) variable by a maximum information entropy principle introduced in [14, 15, 28]. This principle does not require statistical data to define all Fourier modes. It turns out that the asymptotic behavior in time of a weighted variance of the fast modes is sufficient. The necessity of such a strong assumption relies on the fact the gKS equation does not have an infinite-dimensional invariant measure and that we only account for spatial randomness. Via this entropy principle (Theorem 4.1) we then conclude that the Fourier modes of the fast variable  $W$  in (1.9)<sub>2</sub> are Gaussian distributed with zero mean. Hence, we rigorously obtain a noisy gKS equation by applying the random variable  $U = v^\epsilon + W$  in the deterministic equation for the slow variable (1.8). Herewith, our analysis explains how to rigorously add a random force to the gKS equation. Furthermore, our derivation further shows that the induced noise accounts for the unresolved degrees of freedom and hence becomes less important for an increasing number of grid points in computations.

The approach proposed here provides an alternative to the Mori-Zwanzig formalism [26, 40, 41] which depends on Hamiltonian [26, 40] or extended Hamiltonian structures [41]. Mori-Zwanzig techniques and related optimal prediction methods [3] generally rely on a canonical probability distribution (invariant measure) which exists naturally for Hamiltonian systems. The canonical probability density for a Hamiltonian  $H(u)$  is  $\rho(u) := Z^{-1} \exp(-\beta H(u))$ , where  $\beta$  is the inverse temperature and  $Z$  a normalization constant referred to as the partition function. The Mori-Zwanzig formalism then is based on a projection operator  $P$  that projects functions in  $L^2$  onto a subspace that only depends on the resolved degrees of freedom. With respect to the canonical density  $\rho$  such a projection operator  $P$  can be defined by the conditional expectation

$$[Pf](v) := \mathbb{E}[f|v] = \frac{\int f(v, w) \rho(v, w) dw}{\int \rho(v, w) dw}, \quad (1.10)$$

where  $f \in L^2$  and  $v$  is the resolved and  $w$  the unresolved variable. The projection  $P$  and Dyson's formula for evolution operators then provide an equation for the resolved modes  $v$  only. Moreover, (1.10) is the conditional expectation of  $f$  given  $v$  and hence is the best least square approximation of  $f$  by a function of  $v$ . Therefore, the projection  $P$  guarantees optimality which is the key idea in the optimal prediction method. However, neither a Hamiltonian structure nor an invariant measure exists for the gKS equation. Therefore, it is not obvious how to derive standard optimality statements relying on a conditional probability argument [33]. In contrast to such a conditional

probability approach, we achieve optimality in the sense of maximum information entropy. However, we remark that one can also define other projections than (1.10).

The purpose of the present article is threefold: 1. To reliably perform a (stochastic) mode reduction for the full gKS equation in contrast to [33] where a truncated problem is studied. The principal idea is based on an abstract RG approach, as emphasized earlier. We derive error estimates (Theorem 3.6) for this reduction and hence provide rigorous support for the heuristic motivation of a noisy, low dimensional approximation deduced in [4] by the standard RG method in physics [39]; 2. To rigorously support Stinis' assumption of Gaussian distributed Fourier modes [33]. To this end, we derive a probability distribution (Theorem 4.1) for the fast modes by the principle of maximum information entropy. 3. The findings in 1 and 2 form the bases for a new stochastic mode reduction strategy. We are able to reduce the fast variable by an equation for the slow variable only. The information of the fast modes enters as a random variable  $W$  via a force term into the slow mode equations. We are not aware of any previous work that utilizes the RG method in the context of stochastic mode reduction.

We introduce basic notation and well-known results in Section 1.2. A formal derivation of an RG equation for the gKS equation follows in Section 2. In Section 3 we obtain error estimates to rigorously verify the approximation derived in Section 2. In Section 4 we reduce the fast modes by a mode reduction strategy based on the maximum information entropy principle. Finally, in Section 6 we close with conclusions and perspectives.

**1.1. The gKS equation** The KS equation is a paradigmatic model for the study of low-dimensional spatio-temporal chaos or weak/dissipative turbulence as defined by Manneville [23]. This type of turbulence is often characterized by formation of clearly identifiable localized coherent structures in what appears to be a randomly disturbed system, as is e.g. the case with Rayleigh-Bénard convection [31]. The KS equation was first proposed as a model for pattern formation in reaction-diffusion systems by Kuramoto [19]. Its derivation is based on a generalized time-dependent Ginzburg-Landau equation. Sivashinsky [32] derived the KS equation as an asymptotic approximation of a diffusional-thermal flame model. The equation also describes small-amplitude waves on the surface of a thin film flowing down a planar inclined wall (e.g. [12, 18]).

With the addition of the dispersive term,  $u_{xxx}$ , the KS equation becomes the gKS equation. Like the KS equation, it has been reported for a wide variety of systems, from plasma waves with dispersion due to finite ion banana width [5] to a thin film flowing down a planar wall for near-critical conditions (e.g. [29, 18]). The studies in [6, 36, 37] have developed a coherent-structure theory for the interaction of the solitary-pulse solutions of the gKS equation. In [6, 36] the theory was shown to be in agreement with experiments using a thin film coating a vertical fiber, another hydrodynamic system where the gKS equation can be applicable.

The well-posedness of (1.1) is established for example in [35] in the class of generalized Burgers equations which consist of a quadratic nonlinearity and arbitrary linear parabolic part. The article [20] verifies solvability of the gKS equation in bounded domains and studies its limit towards the Korteweg-de Vries equation. In the context of long-time and large-space considerations, there are recent analytical attempts to verify an “equipartition principle” in the power spectrum of periodic solutions by deriving bounds on their space average of  $|u|$  and certain derivatives of it, see [9, 27].

Such a spectral characterization is reminiscent of white noise.

An interesting work that applies the optimal prediction to the KS equation is that of Stinis [33]. Since this approach requires a non-invariant measure, the author constructs a Gibbs measure for the required initial distribution through inference from empirical data (obtained by a computational approach). This allows then to define the conditional expectation providing optimality by an orthogonal projection of the unresolved modes to the resolved ones. However, this approach already assumes a Gaussian distribution from the outset. For this strategy, one also needs to work with the truncated KS equation. Sufficient numerical data are then required in advance for a reliable construction of an initial distribution.

**1.2. Notation** Functions  $u \in H^s(\mathcal{P}_\alpha)$  for  $s \geq 1$  can be represented by their Fourier series,

$$u(x) = \sum_{k \in \mathbb{Z}} u_k \exp\left(i \frac{k}{\alpha} x\right), \quad \overline{u_k} = u_{-k}, \quad (1.11)$$

where  $H^s$  denotes here the usual periodic Sobolev space with finite norm,

$$\|f\|_{H^s}^2 := \sum_{k \in \mathbb{Z}} (1 + |k|^2)^s \left| \hat{f}(k) \right|^2. \quad (1.12)$$

Furthermore, the square root of the latter quantity is a norm on  $H^s(\mathcal{P}_\alpha)$  equivalent to the usual one. We denote for  $s \geq 0$ ,

$$\dot{H}^s(\mathcal{P}_\alpha) := \left\{ u \in H^s(\mathcal{P}_\alpha) \left| \int_{\mathcal{P}_\alpha} u dx = 0 \right. \right\}. \quad (1.13)$$

The subspace of  $\dot{H}^s(\mathcal{P}_\alpha)$  spanned by the set,

$$\left\{ e^{i \frac{k}{\alpha} x} \mid k \in \mathbb{Z}, -N \leq k \leq N \right\} \quad (1.14)$$

is denoted by  $H_N^s$ . For a given integer  $N$  we define the projections  $v := P_N u$  and  $w := Q_N u := (I - P_N)u$  by,

$$\begin{aligned} v = P_N u &= \sum_{|k| \leq N} u_k \exp\left(i \frac{k}{\alpha} x\right), \\ w = Q_N u &= \sum_{|k| > N} u_k \exp\left(i \frac{k}{\alpha} x\right). \end{aligned} \quad (1.15)$$

Let us mention that the gKS equation preserves mass as already noted in Section 1, i.e.,

$$\int_{\mathcal{P}_\alpha} u dx = u_0, \quad (1.16)$$

where  $u_0$  is the zero-th Fourier mode. We remark that  $P_N$  is an orthogonal projection with respect to  $H_N^s$ , that means,

$$\int_{\mathcal{P}_\alpha} (P_N u - u) \phi dx = 0 \quad \text{for all } \phi \in H_N^s. \quad (1.17)$$

The projection  $P_N$  enjoys the following well-known property [13, 22], i.e., for  $k \geq s$ ,  $k \geq 0$  it holds,

$$\|u - P_N u\|_{H^s} \leq C N^{s-k} \|u\|_{H^k} \quad \text{for all } u \in H_N^k(\mathcal{P}_\alpha). \quad (1.18)$$

Next, we introduce Gevrey spaces. For  $\sigma \geq 0$  and  $s \geq 0$  we say that a function  $f$  is in the Gevrey space  $G_{\sigma,s}$  if and only if

$$\|f\|_{G_{\sigma,s}}^2 := \sum_{k \in \mathbb{Z}} \left(1 + |k|^2\right)^s \exp\left(2\sigma\sqrt{1 + |k|^2}\right) |f_k|^2 < \infty, \quad (1.19)$$

where  $f_k$  denote the Fourier coefficients of  $f$ . Note that if  $\sigma = 0$ , then  $H^s = G_{0,s}$ . Moreover, it can be readily be proved, see [17], that for  $u \in G_{\sigma,s}$  the following inequality holds,

$$\|u - P_N u\|_{H^s} \leq N^{s-k} \exp(-\sigma N) \|u\|_{G_{\sigma,k}}. \quad (1.20)$$

**2. Formal derivation of a reduced gKS equation** As noted in Section 1, we adapt RG approaches [24, 25] to the gKS equation.

**2.1. Projections into fast and slow equations** We apply the projections  $P_N$  and  $Q_N$  defined in (1.15) to equation (1.1) and obtain the following coupled system for  $v$  and  $w$ ,

$$\begin{aligned} \partial_t v + P_N B(v + w) + A_v v &= 0, & \text{where } A_v &= P_N A = A P_N, \\ \partial_t w + Q_N B(v + w) + A_w w &= 0, & \text{where } A_w &= Q_N A = A Q_N. \end{aligned} \quad (2.1)$$

We define  $\epsilon = \frac{1}{N^4}$ , where  $N$  is large enough (see error estimates, i.e. Theorem 3.6,) and set,

$$\begin{aligned} \tilde{A}_v &= A_v & \text{on } P_N \dot{H}^s &= H_N^s, \\ \tilde{A}_w &= \epsilon A_w = \frac{A_w}{N^4} & \text{on } Q_N \dot{H}^s &= \dot{H}^s \setminus H_N^s. \end{aligned} \quad (2.2)$$

The eigenvectors of  $\tilde{A}_v$  are the functions  $\exp(i \frac{k}{\alpha} x)$ ,  $k \in \mathbb{Z}$ ,  $|k| \leq N$  with eigenvalues,

$$\rho_k^v := -\nu \left| \frac{k}{\alpha} \right|^2 - i\delta \left( \frac{k}{\alpha} \right)^3 + \kappa \left| \frac{k}{\alpha} \right|^4.$$

Correspondingly, the eigenvectors of  $\tilde{A}_w$  are the functions  $\exp(i \frac{k}{\alpha} x)$ ,  $k \in \mathbb{Z}$ ,  $|k| > N$  with eigenvalues,

$$\rho_k^w := \frac{1}{N^4} \left( -\nu \left| \frac{k}{\alpha} \right|^2 - i\delta \left( \frac{k}{\alpha} \right)^3 + \kappa \left| \frac{k}{\alpha} \right|^4 \right). \quad (2.3)$$

**REMARK 2.1.** *The RG method is formally applied here as if these operators were independent of  $\epsilon$ . This technical step of scaling the linear operator and its subsequent treatment is part of the abstract RG approach introduced in [24, 25] in the context of fluid dynamics.*

We can now rewrite (2.1) by,

$$\begin{aligned} \partial_t v + \tilde{A}_v v + P_N B(v + w) &= 0, \\ \partial_t w + \frac{1}{\epsilon} \tilde{A}_w w + Q_N B(v + w) &= 0. \end{aligned} \quad (2.4)$$

For convenience we additionally define,

$$u = \begin{pmatrix} v \\ w \end{pmatrix}, \quad L = \begin{pmatrix} 0 \\ \tilde{A}_w \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} \tilde{A}_v \\ 0 \end{pmatrix}, \quad F(u) = \begin{pmatrix} -P_N B(v+w) \\ -Q_N B(v+w) \end{pmatrix}, \quad (2.5)$$

and hence rewrite (2.4) in the following compact way,

$$\partial_t u + \frac{1}{\epsilon} L u + \mathcal{A} u = F(u). \quad (2.6)$$

For the subsequent RG analysis we introduce the fast time scale  $s = \frac{t}{\epsilon}$ , and we define  $\tilde{u}(s) = u(\epsilon s)$ . We set  $\tilde{v}(s) = P \tilde{u}(s)$ ,  $\tilde{w}(s) = Q \tilde{u}(s)$ . In this variables (2.4) becomes,

$$\begin{aligned} \partial_s \tilde{v} + \epsilon \tilde{A}_v \tilde{v} + \epsilon P_N B(\tilde{v} + \tilde{w}) &= 0, \\ \partial_s \tilde{w} + \tilde{A}_w \tilde{w} + \epsilon Q_N B(\tilde{v} + \tilde{w}) &= 0, \end{aligned} \quad (2.7)$$

or (2.6),

$$\partial_s \tilde{u} + L \tilde{u} + \epsilon \mathcal{A} \tilde{u} = \epsilon F(\tilde{u}). \quad (2.8)$$

**2.2. Perturbation expansion: The RG equation** We now formally apply the RG method and additionally omit the dependence of  $L$  and  $\mathcal{A}$  on  $N$  as in [24, 25].

We make the ansatz of a naive perturbation expansion,

$$\tilde{u}^\epsilon = \tilde{u}^0 + \epsilon \tilde{u}^1 + \epsilon^2 \tilde{u}^2 + \dots \quad (2.9)$$

for  $\tilde{u}$  in (2.8). After substituting (2.9) into (2.8) we formally obtain the following sequence of problems,

$$\begin{aligned} \partial_s \tilde{u}^0 + L \tilde{u}^0 &= 0, \\ \partial_s \tilde{u}^1 + L \tilde{u}^1 &= F(\tilde{u}^0) - \mathcal{A} \tilde{u}^0, \end{aligned} \quad (2.10)$$

and so on.

Formally, the solution of (2.10)<sub>1</sub> for the initial condition  $\tilde{u}^0(0) = u_0$  is,

$$\tilde{u}^0(s) = \exp(-Ls) u_0. \quad (2.11)$$

Equation (2.11) can be equivalently written by,

$$\begin{aligned} \tilde{v}^0(s) &= v_0, \\ \tilde{w}^0(s) &= \exp(-\tilde{A}_w s) w_0. \end{aligned} \quad (2.12)$$

We solve equation (2.10)<sub>2</sub> with the variation of constants formula,

$$\tilde{u}^1(s) = \exp(-Ls) \int_0^s \exp(L\sigma) [F(\exp(-L\sigma) u_0) - \mathcal{A} \exp(-L\sigma) u_0] d\sigma, \quad (2.13)$$

where  $\tilde{u}^1(0) = 0$ , since we are interested in approximations up to  $O(\epsilon)$  such that  $\tilde{u}^1(0)$  is irrelevant and can be taken to be zero, see [25]. We note that  $\mathcal{A} \exp(-\tilde{A}_w \sigma) = 0$  (since  $P_N Q_N = 0$ ) such that  $\mathcal{A} \exp(-L\sigma) = \mathcal{A}$  and we decompose the rest of the integrand in (2.13) as,

$$\exp(L\sigma) F(\exp(-L\sigma) u_0) - \mathcal{A} v_0 =: F_R(u_0) + \tilde{F}_{NR}(\sigma, u_0), \quad (2.14)$$



where  $F_R(u_0)$  represents the part independent of  $\sigma$  on the left hand side of (2.14) and  $\tilde{F}_{NR}$  the rest. Using standard RG terminology, we refer to  $F_R$  as the “resonant” and  $\tilde{F}_{NR}$  as the “non-resonant” term.

Using (2.11), (2.13), and (2.14) in (2.9) provides the following Duhamel’s form of the formal perturbation expansion for  $\tilde{u} = \tilde{u}^\epsilon$ ,

$$\tilde{u}^\epsilon(s) = \exp(-Ls) \left( u_0 + \epsilon s F_R(u_0) + \epsilon \int_0^s \tilde{F}_{NR}(\sigma, u_0) d\sigma \right) + \mathcal{O}(\epsilon^2). \quad (2.15)$$

The key idea is now to remove the secular term  $\epsilon s F_R(u_0)$  which grows in time. To this end we define the “renormalized function”  $\tilde{U} = \tilde{U}(s)$  as the solution of,

$$\begin{aligned} \partial_s \tilde{U} &= \epsilon F_R(\tilde{U}), \\ \tilde{U}(0) &= u_0. \end{aligned} \quad (2.16)$$

The equation for the slow variable  $U(t) = \tilde{U}(t/\epsilon)$  correspondingly satisfies,

$$\begin{aligned} \partial_t U &= F_R(U), \\ U(0) &= u_0. \end{aligned} \quad (2.17)$$

Let us derive the explicit form of the RG equation for our problem. With the expressions for  $L$  and  $F$ , and the identity  $u_0 = v_0 + w_0$ , we get,

$$\exp(L\sigma)F(\exp(-L\sigma)u_0) = \exp(-L\sigma) \begin{pmatrix} -P_N B(v_0 + \exp(-L\sigma)w_0) \\ -Q_N B(v_0 + \exp(-L\sigma)w_0) \end{pmatrix}. \quad (2.18)$$

Next we identify the resonant terms, i.e.,  $F_R(u_0)$ . With the Fourier series expansion,

$$\phi(x) = \sum_{k \in \mathbb{Z}} \exp\left(i \frac{k}{\alpha} x\right) \phi_k, \quad (2.19)$$

we have,

$$\begin{aligned} B(\phi, \psi) &= i\lambda \sum_{k \in \mathbb{Z}} \exp\left(i \frac{k}{\alpha} x\right) \phi_k \sum_{l \in \mathbb{Z}} \exp\left(i \frac{l}{\alpha} x\right) \frac{l}{\alpha} \psi_l \\ &= i\lambda \sum_{j \in \mathbb{Z}} \exp\left(i \frac{j}{\alpha} x\right) \sum_{k+l=j} \left( \phi_k \frac{j}{\alpha} \right) \psi_l. \end{aligned} \quad (2.20)$$

As a consequence, we end up with the expressions,

$$\begin{aligned} Q_N B(v_0, \exp(-L\sigma)w_0) &= i\lambda \sum_{|j| > N} \exp\left(i \frac{j}{\alpha} x\right) \sum_{\substack{k+l=j \\ |k| \leq N < |l|}} \left( v_{0k} \frac{j}{\alpha} \right) \exp(-\sigma \rho_l^w) w_{0l}, \\ Q_N B(\exp(-L\sigma)w_0, \exp(-L\sigma)w_0) &= i\lambda \sum_{|j| > N} \exp\left(i \frac{j}{\alpha} x\right) \\ &\quad \sum_{\substack{k+l=j \\ |k|, |l| > N}} \left( \exp(-\sigma \rho_k^w) w_{0k} \frac{j}{\alpha} \right) \exp(-\sigma \rho_l^w) w_{0l}, \end{aligned} \quad (2.21)$$

The resonant terms in the first sum are the terms for which  $\rho_l^w = \rho_j^w$  holds. This means,  $\left(-\kappa \left|\frac{l}{\alpha}\right|^2 - i\delta \left(\frac{l}{\alpha}\right)^3 + \kappa \left|\frac{l}{\alpha}\right|^4\right) = \left(-\kappa \left|\frac{j}{\alpha}\right|^2 - i\delta \left(\frac{j}{\alpha}\right)^3 + \kappa \left|\frac{j}{\alpha}\right|^4\right)$ . Since  $\nu, \delta, \kappa > 0$ , the following set characterizes the resonant indices,

$$R_1(j) := \{(k, l) \mid k=0, j=l, |l| > N\}. \quad (2.22)$$

The condition  $\rho_k^w + \rho_l^w = \rho_j^w$  characterizes the resonant terms in the second sum of (2.21), i.e.,  $\left(\frac{k}{\alpha}\right)^n + \left(\frac{l}{\alpha}\right)^n = \left(\frac{j}{\alpha}\right)^n$  for  $n=2, 3, 4$  needs to hold at the same time. Since  $\left|\frac{j}{\alpha}\right|^4 = \left|\frac{k}{\alpha}\right|^4 + \left|\frac{l}{\alpha}\right|^4 + 2\left|\frac{k}{\alpha}\right|^2 \left|\frac{l}{\alpha}\right|^2$ , the set of resonant indices is defined by,

$$R_2(j) := \{(k, l) \mid k=0, l=j, |k|, |l| > N\} \cup \{(k, l) \mid l=0, k=j, |k|, |l| > N\} = \emptyset, \quad (2.23)$$

since  $|k|, |l| > N$ . These considerations determine the resonant part of F by,

$$F_R(u_0) = \begin{bmatrix} -P_N B(v_0) - \tilde{A}_v v_0 \\ -Q_N B_1(v_0, w_0) \end{bmatrix}, \quad (2.24)$$

where  $B_1$  is given by its Fourier series expansions for the corresponding index set  $R_1(j)$ , i.e.,

$$Q_N B_1(v_0, w_0) = 2i\lambda \sum_{|j| > N} e^{i\frac{j}{\alpha}x} \left( v_{00} \frac{j}{\alpha} \right) w_{0j}. \quad (2.25)$$

Equation (2.14) and the above consideration give the non-resonant term by,

$$\tilde{F}_{NR}(\sigma, u_0) = \begin{bmatrix} -P_N B \left( v_0 + e^{-\tilde{A}_w \sigma} w_0, v_0 + e^{-\tilde{A}_w \sigma} w_0 \right) + P_N B(v_0, v_0) + \tilde{A}_v v_0 \\ -Q_N \tilde{B}_1(v_0, w_0) - Q_N \tilde{B}_2(w_0, w_0) \end{bmatrix}, \quad (2.26)$$

where  $Q_N \tilde{B}_1$  and  $Q_N \tilde{B}_2$  are defined by their Fourier series expansions,

$$\begin{aligned} Q_N \tilde{B}_1(v_0, w_0) &= i\lambda \sum_{|j| > N} e^{i\frac{j}{\alpha}x} \sum_{\substack{k+l=j \\ |l| \neq |j| \\ |k| \leq N < |l|}} \left( \left( v_{0k} \frac{j}{\alpha} \right) w_{0l} + \left( w_{0l} \frac{j}{\alpha} \right) v_{0k} \right) e^{(\rho_j^w - \rho_l^w)\sigma}, \\ Q_N \tilde{B}_2(w_0, w_0) &= i\lambda \sum_{|j| > N} e^{i\frac{j}{\alpha}x} \sum_{\substack{k+l=j \\ |k|, |l| > N}} \left( w_{0k} \frac{j}{\alpha} \right) w_{0l} e^{(\rho_j^w - \rho_k^w - \rho_l^w)\sigma}. \end{aligned} \quad (2.27)$$

With (2.24) the RG equation for our problem is in the fast time scale,

$$\begin{aligned} \partial_s \tilde{V} + \epsilon \tilde{A}_v \tilde{V} + \epsilon P_N B(\tilde{V}, \tilde{V}) &= 0, \\ \partial_s \tilde{W} + \epsilon Q_N B_1(\tilde{V}, \tilde{W}) &= 0, \end{aligned} \quad (2.28)$$

or after rescaling by  $t = \epsilon s$ , and denoting  $V = P_N U$ ,  $W = Q_N U$ ,

$$\begin{aligned} \partial_t V + A_v V + P_N B(V, V) &= 0, \\ \partial_t W + Q_N B_1(V, W) &= 0. \end{aligned} \quad (2.29)$$

REMARK 2.2. 1) The above considerations for the resonant and non-resonant terms can easily be extended to general pseudodifferential operators  $P(A)$  with symbol  $p(\xi)$  of the form,

$$\text{Rep}(i\xi) \geq c|\xi|^\nu, \quad |\xi| \rightarrow \infty, \quad (2.30)$$

where  $\nu > 3/2$ . One only needs to adapt the sets for the resonant indices, see (2.22) and (2.23).

2) Note that the  $V$ -equation in the RG equation (2.29) is simply the Galerkin approximation of the gKS equation (1.1).

The special structure of the renormalization equation (2.29)<sub>2</sub> for the unresolved (fast) variable allows to give an explicit expression for its solution. After rewriting (2.29)<sub>2</sub> by

$$\partial_t W_j(t) + 2i\lambda \frac{j}{\alpha} V_0(t) W_j(t) = 0, \quad (2.31)$$

where  $V_0(t) = \text{const.}$  due to conservation of mass (1.16), we immediately obtain the solution,

$$W_j(t) = c_W^j e^{i2\lambda \frac{j}{\alpha} V_0 t}, \quad c_W^j := W_j(0). \quad (2.32)$$

With (2.32) the solution of (2.29)<sub>2</sub> becomes,

$$W(x, t) = \sum_{|j| > N} c_W^j e^{i\frac{j}{\alpha}(x + 2\lambda V_0 t)}. \quad (2.33)$$

Equation (2.33) shows that there is no restriction on the definition of the mass  $V_0$ . In the context of stochastic mode reduction the situation is different, see Section 4.

**2.3. Construction of approximate/renormalized solutions** In order to define renormalized solutions we have first to determine the non-resonant term  $\tilde{F}_{NR}(\sigma, u_0)$  given by (2.26). In fact, we are interested in,

$$F_{NR}(s, U) = \int_0^s \tilde{F}_{NR}(\sigma, U) d\sigma. \quad (2.34)$$

Let

$$\begin{aligned} \text{PF}_{NR}(s, U) &= 2i\lambda \sum_{|j| \leq N} e^{i\frac{j}{\alpha}x} \sum_{\substack{k+l=j \\ |k| \leq N < |l|}} \frac{e^{-\rho_l^w s}}{\rho_l^w} V_k \frac{j}{\alpha} W_l \\ &\quad + i\lambda \sum_{|j| \leq N} e^{i\frac{j}{\alpha}x} \sum_{\substack{k+l=j \\ |k|, |l| > N}} \frac{e^{-(\rho_k^w + \rho_l^w)s}}{\rho_k^w + \rho_l^w} W_k \frac{j}{\alpha} W_l, \\ \text{QF}_{NR}(s, U) &= -2i\lambda \sum_{|j| > N} e^{i\frac{j}{\alpha}x} \sum_{\substack{k+l=j \\ |k| \leq N < |l| \\ |l| \neq |j|}} \frac{e^{(\rho_j^w - \rho_l^w)s} - 1}{\rho_j^w - \rho_l^w} V_k \frac{j}{\alpha} W_l \\ &\quad - i\lambda \sum_{|j| > N} e^{i\frac{j}{\alpha}x} \sum_{\substack{k+l=j \\ |k|, |l| > N}} \frac{e^{(\rho_j^w - \rho_k^w - \rho_l^w)s} - 1}{\rho_j^w - \rho_k^w - \rho_l^w} W_k \frac{j}{\alpha} W_l. \end{aligned} \quad (2.35)$$

Now, we are able to define the approximate solution suggested by the RG theory. We obtain,

$$u^\epsilon(t) = e^{-L\frac{t}{\epsilon}}(U(t) + \epsilon F_{NR}(t/\epsilon, U(t))), \quad (2.36)$$

or with respect to fast  $\overline{w}^\epsilon$  and slow variables  $\overline{v}^\epsilon$ ,

$$\begin{aligned} v^\epsilon &= P_N u^\epsilon = V(t) + \epsilon P F_{NR}(t/\epsilon, U(t)), \\ w^\epsilon &= Q_N u^\epsilon = e^{-Q_N A t} (W(t) + \epsilon Q F_{NR}(t/\epsilon, U(t))). \end{aligned} \quad (2.37)$$

We note that the initial data are defined by,

$$\begin{aligned} v^\epsilon(0) &= V(0) + \epsilon P F_{NR}(0, U(0)) = v_0 + \epsilon P F_{NR}(0, u_0), \\ w^\epsilon(0) &= W(0) + \epsilon Q F_{NR}(0, U(0)) = w_0. \end{aligned} \quad (2.38)$$

**3. The renormalized gKS equation and approximation error** After inserting (2.36) into (1.1) we obtain the following perturbed gKS equation,

$$\partial_t u^\epsilon + A u^\epsilon + B(u^\epsilon, u^\epsilon) = -\epsilon R_\epsilon(U(t)), \quad (3.1)$$

where A and B are defined by (1.3) and  $R_\epsilon$  is given by,

$$\begin{aligned} R_\epsilon &= B\left(e^{-L\frac{t}{\epsilon}}U(t), e^{-L\frac{t}{\epsilon}}F_{NR}(t/\epsilon, U(t))\right) + B\left(e^{-L\frac{t}{\epsilon}}F_{NR}(t/\epsilon, U(t)), e^{-L\frac{t}{\epsilon}}U(t)\right) \\ &\quad + B\left(e^{-L\frac{t}{\epsilon}}F_{NR}(t/\epsilon, U(t)), e^{-L\frac{t}{\epsilon}}F_{NR}(t/\epsilon, U(t))\right) - \epsilon P F_{NR}(t/\epsilon, U(t)) \\ &\quad - e^{-L\frac{t}{\epsilon}}\delta_U F_{NR}(t/\epsilon, U(t))\partial_t U. \end{aligned} \quad (3.2)$$

Next, we study estimates on the approximate solutions  $u^\epsilon$  of equation (3.1). In a first step, we need to investigate the non-resonant part  $F_{NR}$  of the approximate solutions.

**LEMMA 3.1.** *Let  $p \geq 2$  and let  $g \in H^q(\mathcal{P}_\alpha)$  with  $q \geq 4$ . Assume that the solution of the RG equation (2.29) satisfies  $U(t) \in H^p(\mathcal{P}_\alpha)$  for all  $t > 0$ . For  $N$  large enough there exist two constants  $c_1$  and  $C_2$ , where  $C_2$  depends on the initial conditions, and  $c_1$  depends only on  $\mathcal{P}_\alpha$ , but both independent of  $N$ , such that the following estimates are true for all  $t > 0$ ,*

$$\begin{aligned} \|P_N F_{NR}(t/\epsilon, U(t))\|_{H^p} &\leq C_2 e^{-c_1 N^4 t}, \\ \|e^{-Q_N A t} Q_N F_{NR}(t/\epsilon, U(t))\|_{H^p} &\leq C_2 e^{-c_1 N^4 t}. \end{aligned} \quad (3.3)$$

*Proof.*  $c, c_1$ , and  $C_2$  represent generic constants independent of  $N$  (or  $\epsilon$ ). We first derive estimate (3.3)<sub>1</sub>. With the expression (2.35)<sub>1</sub> we immediately obtain,

$$\|P_N F_{NR}(t/\epsilon, U(t))\|_{H^p} \leq c e^{-c N^2 t} (\|V \cdot \nabla W\|_{H^p} + \|W \cdot \nabla V\|_{H^p} + \|W \cdot \nabla W\|_{H^p}), \quad (3.4)$$

where we used the fact that we have the following bound,

$$\frac{e^{-\rho_l^w t/\epsilon}}{\rho_l^w} = \frac{e^{-(\nu|\frac{1}{\alpha}|^2 - i\delta(\frac{1}{\alpha})^3 + \kappa|\frac{1}{\alpha}|^4)t}}{1/N^4(-\nu|\frac{1}{\alpha}|^2 - i\delta(\frac{1}{\alpha})^3 + \kappa|\frac{1}{\alpha}|^4)} \leq c e^{-c N^4 t}. \quad (3.5)$$

The last inequality follows due to  $N < [l]$ . The second estimate  $(3.3)_2$  can be obtained in the same way by using the inequalities  $1 - e^{-x} \leq x$  for all  $x \geq 0$  and  $xe^{-x} \leq \frac{1}{e}$  for all  $x \geq 0$ . We refer the interested reader to [24] for a deeper consideration.  $\square$

The bounds of Lemma 3.1 allow us to control  $R_\epsilon$  in the spirit of [24].

LEMMA 3.2. *For  $N > 0$  and  $g \in H^q(\mathcal{P}_\alpha)$  for  $q \geq 4$ , there exist two constants  $c_1$  and  $C_2$  independent of  $N$ , such that the following estimate holds true for all  $t \geq 0$ ,*

$$\|R_\epsilon(t)\|_{L^2} \leq C_2 e^{-c_1 N^4 t}. \quad (3.6)$$

*Proof.* The proof follows in the same way as the proof of Lemma 3.1. We only need to take into account the expression of  $R_\epsilon$  and apply Lemma 3.1.  $\square$

LEMMA 3.3. *For  $0 < T^* < \infty$  and  $g$  as in Lemma 3.2, there exists an  $0 < \epsilon^* < \infty$  such that for  $0 \leq \epsilon := 1/N^4 \leq \epsilon^*$  solutions to equation (3.1) satisfy  $u^\epsilon \in L^\infty(0, T^*; L^2(\mathcal{P}_\alpha)) \cap L^2(0, T^*; H^2(\mathcal{P}_\alpha))$ .*

*Proof.* The lemma follows by formally testing equation (3.1) with  $u^\epsilon$  and using periodicity of  $\mathcal{P}_\alpha$ , i.e.,  $\frac{\lambda}{6} (\partial_x(u^\epsilon)^3, 1) = 0$ , such that,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u^\epsilon\|^2 + \nu \|\partial_x^2 u^\epsilon\|^2 &\leq \frac{\epsilon^2}{2} \|u^\epsilon\|^2 \\ &+ \frac{1}{2} \|R_\epsilon(U)\|^2 + C(\kappa) \|u^\epsilon\|^2 + \kappa \|\partial_x^2 u^\epsilon\|^2. \end{aligned} \quad (3.7)$$

After defining

$$\begin{aligned} \beta &:= 2(\epsilon^2 + C(\kappa)), \\ \gamma &:= \frac{1}{2} \|R_\epsilon(U)\|^2, \end{aligned} \quad (3.8)$$

we multiply (3.7) by  $\exp\left(-\int_0^t \beta ds\right)$  such that

$$\frac{1}{2} \exp(-\beta t) \frac{d}{dt} \|u^\epsilon\|^2 \leq \exp(-\beta t) \frac{\beta}{2} \|u^\epsilon\|^2 + \exp(-\beta t) \gamma(t). \quad (3.9)$$

Since

$$\frac{d}{dt} \left( \exp(-\beta t) \frac{1}{2} \|u^\epsilon\|^2 \right) = -\beta \exp(-\beta t) \frac{1}{2} \|u^\epsilon\|^2 + \exp(-\beta t) \frac{1}{2} \frac{d}{dt} \|u^\epsilon\|^2, \quad (3.10)$$

we can rewrite (3.9) as

$$\frac{d}{dt} \left( \exp(-\beta t) \frac{1}{2} \|u^\epsilon\|^2 \right) \leq \exp(-\beta t) \gamma(t), \quad (3.11)$$

and subsequent integration together with Lemma 3.2 gives,

$$\|u^\epsilon\|^2(T) \leq C \exp(\beta T) \int_0^T \exp((\beta - C/\epsilon)t) dt \leq \frac{C}{\beta + C/\epsilon} \exp(\beta T). \quad (3.12)$$

For arbitrary  $0 < T^* < \infty$  we can choose  $0 \leq \epsilon \leq \epsilon^* := \frac{1}{\exp(\beta T^*) - \beta/C C_\infty}$  where the constant  $C_\infty$  is chosen such that,

$$\frac{C}{\beta + C/\epsilon} \exp(\beta T^*) \leq C_\infty < C \exp(\beta T^*) / \beta. \quad (3.13)$$

□

The reduced equation for the resolved (slow) modes  $v^\epsilon$  alone follows immediately after using (2.37)<sub>1</sub> and (2.29)<sub>1</sub>, i.e.,  $V(t) = v^\epsilon(t) - \epsilon \text{PF}_{NR}(t/\epsilon, U(t))$ ,

$$\partial_t v^\epsilon + A_v v^\epsilon + P_N B(v^\epsilon, v^\epsilon) = -\epsilon G_\epsilon(U(t), v^\epsilon), \quad (3.14)$$

where the induced force term  $G_\epsilon$  is defined by,

$$\begin{aligned} G_\epsilon(U(t), v^\epsilon) := & P_N B\left(e^{-L_\epsilon^\frac{t}{\epsilon}} v^\epsilon, e^{-L_\epsilon^\frac{t}{\epsilon}} \text{PF}_{NR}(t/\epsilon, U(t))\right) \\ & + P_N B\left(e^{-L_\epsilon^\frac{t}{\epsilon}} \text{PF}_{NR}(t/\epsilon, U(t)), e^{-L_\epsilon^\frac{t}{\epsilon}} v^\epsilon\right) \\ & + P_N B\left(e^{-L_\epsilon^\frac{t}{\epsilon}} \text{PF}_{NR}(t/\epsilon, U(t)), e^{-L_\epsilon^\frac{t}{\epsilon}} \text{PF}_{NR}(t/\epsilon, U(t))\right) \\ & - A_v \text{PF}_{NR}(t/\epsilon, U(t)) - e^{-L_\epsilon^\frac{t}{\epsilon}} \delta_U \text{PF}_{NR}(t/\epsilon, U(t)) \partial_t U, \end{aligned} \quad (3.15)$$

and  $U(t) = V(t) + W(t) = v^\epsilon(t) + W(t)$  is the solution of the RG equation (2.29).

LEMMA 3.4. *For  $0 < T^* < \infty$  and  $g$  as in Lemma 3.2, there exists an  $0 \leq \epsilon^* < \infty$  such that for  $0 \leq \epsilon = \frac{1}{N^4} \leq \epsilon^*$  solutions to (3.14) satisfy  $v^\epsilon := P_N u^\epsilon \in L^\infty(0, T^*; L^2(\mathcal{P}_\alpha)) \cap L^2(0, T^*; H^2(\mathcal{P}_\alpha))$ .*

*Proof.* The proof is similar to the proof of Lemma 3.3. □

The same arguments as those for Lemma 3.2 lead to the following.

LEMMA 3.5. *For  $N > 0$  and  $v_0 = P_N g \in H^q(\mathcal{P}_\alpha)$  for  $q \geq 4$ , there exist two constants  $c_1$  and  $C_2$  independent of  $N$ , such that the following estimate holds true for all  $t \geq 0$ ,*

$$\|G_\epsilon(t)\|_{L^2} \leq C_2 e^{-c_1 N^4 t}. \quad (3.16)$$

Estimate (3.16) shows that the RG method provides an invariant manifold by the solutions  $v^\epsilon$  such that all orbits are exponentially attracted. The following theorem gives qualitative information about the RG approach by quantifying the error between (3.14) and (1.1).

THEOREM 3.6. *Let  $g \in H^4(\mathcal{P}_\alpha)$ ,  $\epsilon = \frac{1}{N^4}$ , and suppose that  $u, u^\epsilon \in L^\infty(0, T; H^2(\mathcal{P}_\alpha))$ . Then, the difference between the reduced solution  $v^\epsilon$  and the exact solution of the gKS equation (1.1) satisfies the following error estimate,*

$$\|u - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)}^2(T) \leq C\epsilon^2 + \exp(CT) \left( \epsilon^{1/4} + \epsilon \right). \quad (3.17)$$

*If we suppose that  $u, u^\epsilon \in L^\infty(0, T; G_{\sigma,2}(\mathcal{P}_\alpha))$ , then we can improve (3.17) in the following way,*

$$\|u - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)}^2(T) \leq C\epsilon^2 + \exp(CT) \left( \epsilon^{1/4} \exp\left(-\frac{\sigma}{\epsilon^{1/4}}\right) + \epsilon \right). \quad (3.18)$$

REMARK 3.7. *The exponential growth in time is not surprising, see for example estimate (3.12) in the proof of Lemma 3.3. This estimate motivates the definition of a new variable  $h(x, t; \eta) := \exp(-\eta t)u(x, t)$ . Tadmor verifies in [35] global existence for such a decayed variable  $h$  and a conservative form of the KS equation.*

*Proof.* The error  $\|u - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)}$  can be bounded using the triangle inequality by,

$$\|u - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)} \leq \|u - u^\epsilon\|_{L^2(\mathcal{P}_\alpha)} + \|u^\epsilon - v^\epsilon\|_{L^2(\mathcal{P}_\alpha)}, \quad (3.19)$$

where the first term on the right-hand side in (3.19) represents the approximation error from the RG method (*RG error*) and the second term accounts for the truncation error (*Tr error*). For notational brevity, we introduce the error variables

$$e_{RG}^\epsilon := u - u^\epsilon, \quad \text{and} \quad e_{Tr}^\epsilon := u^\epsilon - v^\epsilon. \quad (3.20)$$

**Step 1: (RG error)** The equation for the error variable  $e_{RG}^\epsilon$  reads,

$$\partial_t e_{RG}^\epsilon + [\kappa \partial_x^2 + \delta \partial_x^3 + \nu \partial_x^4] e_{RG}^\epsilon + e_{RG}^\epsilon \partial_x u + u^\epsilon \partial_x e_{RG}^\epsilon = \epsilon R_\epsilon(U). \quad (3.21)$$

First, we test (3.21) with  $-\partial_x^2 e_{RG}^\epsilon$ , i.e.,

$$\begin{aligned} \partial_t (\partial_x e_{RG}^\epsilon, \partial_x e_{RG}^\epsilon) - (\kappa \partial_x^2 e_{RG}^\epsilon, \partial_x^2 e_{RG}^\epsilon) - (\delta \partial_x^3 e_{RG}^\epsilon, \partial_x^2 e_{RG}^\epsilon) - (\nu \partial_x^4 e_{RG}^\epsilon, \partial_x^2 e_{RG}^\epsilon) \\ - (e_{RG}^\epsilon \partial_x u, \partial_x^2 e_{RG}^\epsilon) - (u^\epsilon \partial_x e_{RG}^\epsilon, \partial_x^2 e_{RG}^\epsilon) = -(\epsilon R_\epsilon(U), \partial_x^2 e_{RG}^\epsilon). \end{aligned} \quad (3.22)$$

Then, we use the test function  $e_{RG}^\epsilon$ ,

$$\begin{aligned} \partial_t (e_{RG}^\epsilon, e_{RG}^\epsilon) + (\kappa \partial_x^2 e_{RG}^\epsilon, e_{RG}^\epsilon) + (\delta \partial_x^3 e_{RG}^\epsilon, e_{RG}^\epsilon) + (\nu \partial_x^4 e_{RG}^\epsilon, e_{RG}^\epsilon) \\ + (e_{RG}^\epsilon \partial_x u, e_{RG}^\epsilon) + (u^\epsilon \partial_x e_{RG}^\epsilon, e_{RG}^\epsilon) = (\epsilon R_\epsilon(U), e_{RG}^\epsilon). \end{aligned} \quad (3.23)$$

Next, we add up (3.22) and (3.23) and apply the Sobolev embedding theorem and standard inequalities to end up with,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} [\|e_{RG}^\epsilon\|^2 + \|\partial_x e_{RG}^\epsilon\|^2] + (\nu - 3\alpha) [\|\partial_x^2 e_{RG}^\epsilon\|^2 + \|\partial_x^3 e_{RG}^\epsilon\|^2] \\ \leq C_{RG}(\kappa, \alpha, \epsilon, \|u\|_{H^1}, \|\partial_x u\|_{H^1}, \|u^\epsilon\|_{H^1}, \|\partial_x u^\epsilon\|_{H^1}) \|e_{RG}^\epsilon\|_{H^1}^2. \end{aligned} \quad (3.24)$$

After defining

$$\tilde{C}_{RG} = 2C_{RG}, \quad (3.25)$$

we can multiply (3.24) by  $\exp\left(-\int_0^t \tilde{C}_{RG} ds\right)$  such that

$$\frac{1}{2} \exp\left(-\int_0^t \tilde{C}_{RG} ds\right) \frac{d}{dt} \|e_{RG}^\epsilon\|^2 \leq \exp\left(-\int_0^t \tilde{C}_{RG} ds\right) \frac{\tilde{C}_{RG}}{2} \|e_{RG}^\epsilon\|^2. \quad (3.26)$$

Applying a corresponding identity based on the product rule as (3.10) in the proof of Lemma 3.3, we can simplify (3.26) to,

$$\frac{d}{dt} \left( \exp\left(-\int_0^t \tilde{C}_{RG} ds\right) \frac{1}{2} \|e_{RG}^\epsilon\|^2 \right) \leq 0, \quad (3.27)$$

which further reduces by assumptions of Theorem 3.6 and after integration to,

$$\frac{1}{2} \|e_{RG}^\epsilon\|^2(T) \leq \frac{1}{2} \exp\left(\int_0^T \tilde{C}_{RG} ds\right) \|e_{RG}^\epsilon\|^2(0) \leq C \|e_{RG}^\epsilon\|^2(0). \quad (3.28)$$

In order to get a bound controlled by  $\epsilon$  on the right-hand side of (3.28), we have to take into account the definition of the initial data (2.38), i.e.,

$$\|e_{RG}^\epsilon\|^2(0) = \epsilon^2 \|\text{PF}_{NR}(0, g)\|^2 \leq C\epsilon^2, \quad (3.29)$$

since  $g \in H^2(\mathcal{P}_\alpha)$  and hence  $\text{PF}_{NR}(0, g) \in H^1(\mathcal{P}_\alpha)$  and its norm is bounded independently of  $\epsilon = \frac{1}{N^4}$  in  $H^1(\mathcal{P}_\alpha)$ . Hence, we can conclude that,

$$\|e_{RG}^\epsilon\|_{L^2(\mathcal{P}_\alpha)}^2(t) \leq C\epsilon^2, \quad (3.30)$$

which holds uniformly in time.

**Step 2: (Tr error)** We derive an estimate for the error variable  $e_{Tr}^\epsilon := u^\epsilon - v^\epsilon = u^\epsilon - \text{P}_N u^\epsilon$ . From (3.1) and (3.14), the error  $e_{Tr}^\epsilon$  satisfies the equation,

$$\partial_t e_{Tr}^\epsilon + [\kappa \partial_x^2 + \delta \partial_x^3 + \nu \partial_x^4] e_{Tr}^\epsilon + \text{P}_N[u^\epsilon u_x^\epsilon] - v^\epsilon v_x^\epsilon = \epsilon(\text{P}_N R_\epsilon(U) - G_\epsilon(U, v^\epsilon), v^\epsilon), \quad (3.31)$$

which can be rewritten for all  $\phi \in H_N^2$  by

$$\begin{aligned} \partial_t(e_{Tr}^\epsilon, \phi) + ([\kappa \partial_x^2 + \delta \partial_x^3 + \nu \partial_x^4] e_{Tr}^\epsilon, \phi) \\ + (\{\text{P}_N[u^\epsilon u_x^\epsilon] - v^\epsilon v_x^\epsilon\}, e_{Tr}^\epsilon) = \epsilon(\{\text{P}_N R_\epsilon(U) - G_\epsilon(U, v^\epsilon)\}, \phi). \end{aligned} \quad (3.32)$$

Choosing  $\phi = e_{Tr}^\epsilon$  allows to estimate (3.32) in the following way,

$$\frac{1}{2} \frac{d}{dt} \|e_{Tr}^\epsilon\|^2 + \nu \|\partial_x^2 e_{Tr}^\epsilon\|^2 \leq C(\alpha, \kappa) \|e_{Tr}^\epsilon\|^2 + \alpha \|\partial_x e_{Tr}^\epsilon\|^2 + (I) + (II), \quad (3.33)$$

where we define,

$$\begin{aligned} (I) &:= (\text{P}_N[u^\epsilon u_x^\epsilon] - v v_x, e_{Tr}^\epsilon), \\ (II) &:= \epsilon(\|\text{P}_N R_\epsilon(U)\| + \|G_\epsilon(U, v^\epsilon)\|) \|e_{Tr}^\epsilon\|. \end{aligned} \quad (3.34)$$

Let us first control term (I), that means,

$$\begin{aligned} |(I)| &\leq |(-\text{P}_N u^\epsilon (\text{P}_N u^\epsilon)_x + u^\epsilon u_x^\epsilon, e_{Tr}^\epsilon)| + |(\text{P}_N u^\epsilon (\text{P}_N u^\epsilon)_x - v^\epsilon v_x^\epsilon, e_{Tr}^\epsilon)| \\ &\leq \frac{1}{2} |(\partial_x((u^\epsilon - \text{P}_N u^\epsilon)(u^\epsilon + \text{P}_N u^\epsilon)), e_{Tr}^\epsilon)| + 1/4 |(\partial_x(\text{P}_N u^\epsilon + v^\epsilon), e_{Tr}^\epsilon)| \\ &\leq C(\|u^\epsilon\|_{H^1} + \|\text{P}_N u^\epsilon\|_{H^1}) \|u^\epsilon - \text{P}_N u^\epsilon\|_{H^1} \|e_{Tr}^\epsilon\| + C \|e_{Tr}^\epsilon\|^2, \end{aligned} \quad (3.35)$$

where we used the embedding  $H_N^1(\mathcal{P}_\alpha)$  into  $L^\infty(\mathcal{P}_\alpha)$ , i.e.,  $\|\partial_x(\text{P}_N u^\epsilon + v^\epsilon)\|_{L^\infty} \leq C$ . The second term (II) immediately becomes,

$$|(II)| \leq \epsilon \exp(-ct/\epsilon) C \|e_{Tr}^\epsilon\|. \quad (3.36)$$

Define

$$\begin{aligned} \gamma(t) &:= C(\|u^\epsilon\|, \|v^\epsilon\|, \|u^\epsilon\|_{H^1}, \|\text{P}_N u^\epsilon\|_{H^1}, \|u^\epsilon\|_{H^2}) \epsilon^{1/4} \\ &\quad + C(\|u^\epsilon\|, \|v^\epsilon\|) \epsilon \exp(-Ct/\epsilon), \\ \beta &:= 2(C(\alpha, \kappa) + C), \end{aligned} \quad (3.37)$$

and mulitply (3.33) with  $\exp\left(-\int_0^t \beta ds\right)$  such that

$$\frac{1}{2} \exp(-\beta t) \frac{d}{dt} \|e_{Tr}^\epsilon\|^2 \leq \frac{\beta}{2} \exp(-\beta t) \|e_{Tr}^\epsilon\|^2 + \gamma(t) \exp(-\beta t). \quad (3.38)$$



Using again a corresponding identity to (3.10) in Lemma 3.3 we can rewrite (3.38) as,

$$\frac{d}{dt} \left( \exp(-\beta t) \frac{1}{2} \|e_{Tr}^\epsilon\|^2 \right) \leq \gamma(t) \exp(-\beta t), \quad (3.39)$$

which becomes after integration with respect to time,

$$\begin{aligned} \frac{1}{2} \exp(-\beta T) \|e_{Tr}^\epsilon\|^2(T) &\leq C \int_0^T \left\{ \left( \epsilon^{1/4} + \epsilon \exp(-Ct/\epsilon) \right) \exp(-\beta t) \right\} dt \\ &\leq \frac{C\epsilon^{1/4}}{\beta} (1 - \exp(\beta T)) + \frac{C\epsilon}{\beta + C/\epsilon} (1 - \exp(-(\beta + C/\epsilon)T)). \end{aligned} \quad (3.40)$$

In the remaining part we want to improve (3.40) with the help of Gevrey spaces. To this end, we remark that the factor  $\epsilon^{1/4}$  in (3.37) relies on the interpolation estimate (1.18). If we assume that solutions  $u^\epsilon$  are in  $G_{\sigma,s}$ , then we improve (1.18) by (1.20). As a consequence, we are able to rewrite inequality (3.40) by

$$\begin{aligned} \frac{1}{2} \|e_{Tr}^\epsilon\|^2(T) &\leq C \int_0^T \left( \epsilon^{1/4} \exp\left(\beta(T-t) - \frac{\sigma}{\epsilon^{1/4}}\right) + \epsilon \exp(-Ct/\epsilon + \beta(T-t)) \right) dt \\ &\leq \frac{C\epsilon^{1/4}}{\beta} \exp(\beta T) + \frac{C\epsilon}{\beta + c/\epsilon} \exp(\beta T). \end{aligned} \quad (3.41)$$

□

**4. Stochastic mode reduction** In this section the renormalized equations (2.29), (2.31) from Sections 2.1 - 3 allow for a rigorous stochastic mode reduction similar in spirit to the Mori-Zwanzig one [26, 40] but for systems not satisfying an extended or generalized Hamiltonian structure [41] and without an intrinsic invariant measure, as is required for Mori-Zwanzig.

As in the Mori-Zwanzig formalism, we assign a stochastic process to the unresolved modes. This is done by applying Jaynes maximum entropy principle, see [15, 14, 28]. This seems a reasonable approach for our problem since we do not have a canonically induced probability density. Hence maximizing the information entropy for the probability density of Fourier modes is equivalent to maximizing the multiplicity of Fourier modes. Multiplicity means the number of different ways a certain state in a system can be achieved. States in a system with the highest multiplicity can be realized by nature in the largest number of ways. Hence, the probability density functions with maximum entropy are optimal statistical descriptions.

It should also be noted that a system at equilibrium will most probably be found in the state of highest multiplicity since fluctuations from that state will be usually too small to measure. The probability distribution may also be obtained from experiments as statistical data. In [33], the probability distribution is constructed by a conditional expectation obtained from previously computed samples which are used to fit an a priori assumed Gaussian distribution.

Finally, we emphasize that the maximum entropy principle can also be applied to problems where one lacks deterministic data as a consequence of not enough experimental data to fix all degrees of freedom. A common approach to model such uncertainty is to use white noise. The maximum entropy method turns out to be an attractive alternative because we do not have to assume the probability distribution.

However, since we apply the entropy maximization principle [15] on an approximate equation, we already neglect information from the beginning and hence have to account for this by an asymptotic in time characterization of the fast modes for example, see Assumption (A) below. This assumption might be improved or adapted appropriately in other applications.

**4.1. Problem induced probability density by maximizing information entropy** With the considerations at the beginning of Section 4, we assign a probability distribution to the unresolved degrees of freedom  $W$  based on the following

**Assumptions:**

- (A) For the unresolved degrees of freedom  $W$  we define the *weighted variance*  $\mathbb{E}[\mathcal{V}(W)]$  by,

$$\mathbb{E}[\mathcal{V}(W)] := \mathbb{E} \left[ \frac{12}{\pi^2} \sum_{|k| > N} \left( k^{-3/2} W_k \right)^2 \right], \quad (4.1)$$

for arbitrary  $\sigma_\infty^2 > 0$  and  $\beta \in \mathbb{R}$ . We assume that the fast modes already asymptotically reached their equilibrium state in time, that means,

$$\mathbb{E}[\partial_t \mathcal{V}(W)] = \sigma_\infty^2 t^{\beta-1}, \quad (4.2)$$

for all  $t > 0$  large enough.

- (B) Under (i) the probability density function  $P[W_{N+1} \leq w_{N+1}, W_{N+2} \leq w_{N+2}, \dots] := F(w_{N+1}, w_{N+2}, \dots)$ , i.e.,

$$F(w_{N+1}, w_{N+2}, \dots) := \int_{-\infty}^{w_{N+1}} \int_{-\infty}^{w_{N+2}} \dots f(w_{N+1}, w_{N+2}, \dots) dw_{N+1} dw_{N+2} \dots, \quad (4.3)$$

has maximum *information entropy*  $\mathcal{S}_I(f)$ ,

$$\mathcal{S}_I(f) = - \int f(\mathbf{w}) \log \left( \frac{f(\mathbf{w})}{\nu(\mathbf{w})} \right) d\mathbf{w}, \quad (4.4)$$

where  $d\mathbf{w} := dw_{N+1} dw_{N+2} \dots$ ,  $f(\mathbf{w})$  denotes the probability density of the unresolved variable  $W$  and  $\nu(\mathbf{w})$  is an *invariant measure* which is defined on background information intrinsically given by the physical origin of  $W$ .

REMARK 4.1. 1) The idea of deriving probability distributions for multiscale evolution problems by maximizing the information entropy seems to go back to [21]. The energy argument in [21], which assumes that the fast modes reached already the stationary state, does not provide here enough information to fix the Lagrange multiplier  $\lambda_1$  associated with this energy constraint. We impose Assumption (A) instead.

2) A mechanical system governed by the Hamiltonian  $H(q, p)$  canonically induces an invariant measure by the density distribution function  $f(q, p) := \frac{1}{Z(\beta)} e^{-\beta H(q, p)}$ .

3) Equation (4.2) accounts for the fact that the fast modes do not entirely contribute to an invariant measure. For simplicity, we also neglect a possible random contribution to the evolution in time. This is a further reason for the strong assumption in (4.2).

In information theory, an entropy related to (4.4) was originally introduced by Shannon [30] to measure the maximum information content in a message. The Assumptions (A) and (B) above account for the lack of a free energy and a Hamiltonian

for which the thermodynamic equilibrium (invariant measure) can be achieved via the gradient flow with respect to the Wasserstein distance [16]. In fact, it should be noted that minimizing the free energy with respect to constant internal energy is equivalent to maximizing the entropy.

**THEOREM 4.1.** *Under Assumptions **(A)** and **(B)**, it follows that the unresolved modes  $W_k$  for  $|k| > N$  obtained by equation (2.31) are normally distributed with zero mean, i.e.,  $\mu_k = 0$ , and variance  $\sigma_k = \frac{k^2 \alpha}{2\lambda_1 V_0}$ , where  $\lambda_1(t) := \frac{t^{1-\beta}}{\sigma_\infty^2}$  is a Lagrange multiplier.*

To keep the considerations simple, we only account for a spatial random process and keep the time deterministic in Theorem 4.1 (and Assumption **(A)**).

*Proof.* To maximize (4.4) under Assumptions **(A)** and **(B)**, we apply the following constraints:

$$(CI) \quad \begin{cases} \int f(\mathbf{w}) d\mathbf{w} = 1, \\ E[\partial_t \mathcal{V}(W)] := \int f(\mathbf{w}) \partial_t \mathcal{V}(\mathbf{w}) d\mathbf{w} = \sigma_\infty^2 t^{\beta-1}, \end{cases} \quad (4.5)$$

where (4.5)<sub>2</sub> is a consequence of assumption (i).

Since a Hamiltonian structure for our problem does not exist as already noted, it is not clear how to define the invariant measure  $\nu$  under the constraints (4.5). We will see in equations (4.14) and (4.17) that the following simple choice for  $\nu$  will be sufficient,

$$\nu(\mathbf{w}) := \prod_{|k| > N} \frac{1}{c_{W_k}}. \quad (4.6)$$

**REMARK 4.2.** *The measure  $\nu(\mathbf{w})$  is the probability density function if we only have a priori information. Usually, it is a nontrivial task and basic considerations of symmetries are required to find this measure  $\nu$ .*

Hence, maximizing the entropy  $S_I$  subject to the constraints (4.5) leads to

$$\int \delta f(\mathbf{w}) \left\{ \log \left( \frac{f(\mathbf{w})}{\nu(\mathbf{w})} \right) + \nu(\mathbf{w}) + \lambda_0 + \lambda_1 \partial_t \mathcal{V}(\mathbf{w}) \right\} d\mathbf{w} = 0, \quad (4.7)$$

where  $\lambda_\ell$  for  $\ell = 0, 1$  are Lagrange multipliers associated with the constraints (4.5). In order to give (4.5)<sub>2</sub> a precise meaning, we write down the explicit form of the equation belonging to each Fourier coefficient of the fast mode variable  $w = Q_N u$  solving (2.29).

We briefly show what the constraint (4.5)<sub>2</sub> means,

$$\begin{aligned} E[\partial_t \mathcal{V}(W)] &= \frac{12}{\pi^2} \int f(\mathbf{w}) \sum_{|k| > N} w_k \partial_t w_k d\mathbf{w} \\ &= \frac{12}{\pi^2} \sum_{|k| > N} \frac{V_0}{k^2 \alpha} \int f(\mathbf{w}) w_k^2 d\mathbf{w} = \sigma_\infty^2 t^{\beta-1}. \end{aligned} \quad (4.8)$$

**REMARK 4.3.** *If we suppose zero mass, then we obtain for the variable  $V_0$  representing the zeroth Fourier mode in (2.31), (2.32), and (2.33) that  $V_0 = 0$ . As a consequence, the solution of the renormalization equation of the unresolved (fast) variable  $W(x, t)$  is defined by its initial condition in this case,*

$$W(x) = w_0(x), \quad (4.9)$$

see (2.31). This raises the question of whether it might be not more suitable to generally work in the case of mass  $V_0 = 1$ ?

Since (4.7) should hold for arbitrary variations  $\delta f$ , we obtain the following expression for the probability density function,

$$f(\mathbf{w}) = \frac{1}{Z} \nu(\mathbf{w}) e^{-\lambda_1 \partial_t \mathcal{V}(\mathbf{w})}, \quad (4.10)$$

where  $Z := e^{\nu(\mathbf{w}) + \lambda_0}$  is called the “partition function” which is determined by the normalization constraint (4.5)<sub>2</sub>, i.e.,

$$Z := Z(\lambda_1) = \int \nu(\mathbf{w}) e^{-\lambda_1 \partial_t \mathcal{V}(\mathbf{w})} d\mathbf{w}. \quad (4.11)$$

Since the constraint (4.5)<sub>2</sub> is quadratic in its nature, we represent it by

$$\lambda_1 \partial_t \mathcal{V}(\mathbf{w}) = \frac{12}{\pi^2} \sum_{|k| > N} \frac{\lambda_1 V_0}{k^2 \alpha} w_k^2 = \frac{12}{\pi^2} \sum_{|k| > N} \frac{1}{k^2} \frac{1}{2\sigma_W^2} (w_k - \mu_k)^2, \quad (4.12)$$

where

$$\sigma_W^2 = \frac{\alpha}{2\lambda_1 V_0}, \quad \text{and} \quad \mu_k = 0. \quad (4.13)$$

With identities (4.12) and (4.13) the probability density function (4.10) can be written as

$$f(\mathbf{w}) = \frac{1}{Z} \prod_{|k| > N} c_{W_k}^{-1} k \sigma_W \sqrt{2\pi} \mathcal{N}(\mu_k, k \sigma_W, w_k), \quad (4.14)$$

where  $\mathcal{N}$  is the normal distribution given by

$$\mathcal{N}(\mu_k, k \sigma_W, w_k) = \frac{1}{k \sigma_W \sqrt{2\pi}} e^{-\frac{(w_k - \mu_k)^2}{2k^2 \sigma_W^2}}, \quad (4.15)$$

which is characterized by the following moments

$$\begin{aligned} \int_{-\infty}^{\infty} \mathcal{N}(\mu, \sigma, w) dw &= 1, \quad \int_{-\infty}^{\infty} \mathcal{N}(\mu, \sigma, w) w dw = \mu, \quad \text{and} \\ \int_{-\infty}^{\infty} \mathcal{N}(\mu, \sigma, w) w^2 dw &= \sigma^2 + \mu^2. \end{aligned} \quad (4.16)$$

The first property in (4.16) together with the normalization condition (4.5)<sub>1</sub> allow us to define the partition function  $Z$  by

$$Z = \prod_{|k| > N} c_{W_k}^{-1} k \sigma_W \sqrt{2\pi}. \quad (4.17)$$

The probability density function  $f$  admits then the simple form as a product of Gaussian distributions, i.e.,

$$f(\mathbf{w}) = \prod_{|k| > N} \mathcal{N}(\mu_k, k \sigma_W, w_k). \quad (4.18)$$

With the second and third property in (4.16) and the constraint (4.5)<sub>2</sub>, i.e., (4.8), we obtain for all  $|k| > N$ ,

$$\sigma_\infty^2 t^{\beta-1} = \mathbb{E}[\partial_t \mathcal{V}(W)] = \frac{12}{\pi^2} \sum_{|k| > N} \frac{V_0}{k^2 \alpha} \sigma_W^2 = \frac{12}{\pi^2} \sum_{|k| > N} \frac{1}{2k^2 \lambda_1}. \quad (4.19)$$

We conclude with (4.19) that the Lagrange parameter  $\lambda_1$  is,

$$\lambda_1 = \frac{12}{\pi^2} \sum_{|k| > N} \frac{1}{2k^2 \sigma_\infty^2} = \frac{t^{1-\beta}}{\sigma_\infty^2}. \quad (4.20)$$

The important information contained in formula (4.20) and (4.13)<sub>1</sub> is that we do not have to assert to each Fourier mode  $k$  its standard deviation  $k^2 \sigma_W^2$ . We only need to determine once the Lagrange parameter  $\lambda_1$  via (4.20). We already know from the beginning ((4.19)) how to choose the mean, i.e.,  $\mu_k = 0$  for all  $|k| > N$ .  $\square$

Hence, the approach of maximizing the generalized information entropy allows to systematically determine the probability distribution function  $f(\mathbf{w})$  of the Fourier modes for the unresolved degrees of freedom  $W$ .

We emphasize that the RG approach suggests multiplicative noise as a compensation for the unresolved modes unlike the commonly obtained additive noise by Mori-Zwanzig's mode reduction [41]. Moreover, an estimate (3.6), which can be correspondingly derived by additionally accounting for the Galerkin error, shows that the influence of the stochastic force decreases for decreasing  $\epsilon := \frac{1}{N^4}$ .

**5. Direct approach: Replacement of  $G_\epsilon$  by white noise** The result of Lemma 3.5 also enables for a direct approach to model the unresolved degrees of freedom as completely unknown. Such kind of complete uncertainty is generally described by white noise  $W(x)$  with zero mean and a variance equal to the power spectral density. It is very common and widely accepted to model uncertainty by white noise. Hence, we replace  $\epsilon G_\epsilon(U(t), v^\epsilon)$  in equation (3.14) by

$$\mathcal{N}_\epsilon(x, t) := \epsilon \exp(-Ct/\epsilon) W(x), \quad (5.1)$$

where  $W(x) \in L^2(\mathcal{P}_\alpha)$  is the Gaussian random variable as motivated above, i.e., with zero mean  $\mu$  and suitable variance  $\sigma$ . It is immediately clear that  $\mathcal{N}_\epsilon$  is a compatible replacement of  $G_\epsilon$  since (5.1) satisfies a bound corresponding to the one in Lemma 3.5. One can follow Stinis' approach [33] for example in order to determine  $\mu$  and  $\sigma$  by a maximum likelihood method.

**6. Discussion and conclusions** We have formally developed a new stochastic mode reduction strategy with a rigorous basis by obtaining appropriate error estimates. The analysis can be summarized in three key steps as follows:

(1) *RG method:* The RG technique [24, 25] turns out to be a formal and feasible method to decompose the gKS equation into slow  $v^\epsilon$  and fast variables  $w^\epsilon$ , respectively. The equation for the slow modes  $v^\epsilon$  represents a Galerkin approximation of the gKS equation plus an additional perturbed force term  $\epsilon G_\epsilon$  which also depends on the infinite dimensional renormalized fast modes  $W$ . An important property of the RG technique is that it can be easily extended to higher space dimensions, see [24]. We also remark that the dispersion term, i.e.  $u_{xxx}$ , does not affect the mode reduction analysis.

(2) *Error bounds:* We rigorously characterize the formal RG method (1) by qualitative error estimates (Theorem 3.6). These estimates further allow for an additional direct mode reduction strategy which is much simpler and straightforward but not as systematic. The basic idea is to replace the perturbed force term  $\epsilon G_\epsilon$  directly by white noise. A physical motivation for such a simplified reduction is the fact that white noise is a well-accepted random model for complete uncertainty.

(3) *Maximum entropy principle:* Due to the lack of a Hamiltonian structure and an invariant measure, we apply Jaynes' maximum entropy principle [15, 14] to define the renormalized fast modes  $W$  as a random variable. This random variable then, together with the renormalized approximation of the slow variable  $v^\epsilon$ , provides a systematic explanation for the appearance of a noisy low dimensional gKS equation. In contrast to optimal prediction we obtain optimality in the sense of maximum entropy here.

There are three main features of the new low dimensional gKS equations:

(i) *Reliable and efficient numerics:* The low dimensional formulation developed here should allow for reliable (since information from the unresolved degrees of freedom included) and efficient (since low dimensional) numerical approximations. In fact, we systematically account for the unresolved degrees of freedom by the steps (1) and (2) above. This is especially of importance since the choice of slow and fast variables depends on the physical problem and is often not clear. For instance, by considering the gKS in large domains, it is possible to introduce a further scale which accounts for the unstable modes. Hence, we can study three different scales such as "unstable modes", "slow stable modes", and "fast stable modes". The main question is then how to account for the unstable and the fast stable modes in an equation for the resolved slow modes only.

Moreover, the error estimates from step (2) provide a qualitative measure on how to choose the dimension of the slow variable. This is also the main advantage of mode reduction considerations over pure convergence analyses of Galerkin approximations (e.g. numerical schemes) where one completely neglects the unresolved degrees of freedom. Hence, straightforward discretization strategies might lose model relevant information in the neglected degrees of freedom. This is a major motivation to include rigorous mode reduction strategies as an important part of the development of computational schemes. We further remark that this is a major reason why the application of noise to deterministic partial differential equations shows good results and is currently a topic of increasing interest.

(ii) *No Hamiltonian structure; no invariant measure:* Many classical mode reduction strategies rely either on a Hamiltonian structure or an invariant measure. Based on the three steps (1)-(3) above, the new asymptotic reduction strategy circumvents such dependences. For example, when classical optimal prediction methods [3] fail because of such deficiencies, the stochastic renormalization provides optimality in the sense of maximum information entropy and hence proves as a promising alternative.

(iii) *The role of noise:* We gain a rigorous understanding of the origin of noise and the way it appears in the gKS equation. This is especially of interest due to numerical evidence provided together with a heuristic motivation in [4] for instance.

Clearly, there are open questions and future perspectives. For example, motivated by the comparative study initiated by Stinis [34], it would be of interest to numerically analyze and compare available mode reduction strategies such as adiabatic elimination

[38], invariant manifolds [7], and optimal prediction [3] with the new RG approach developed here. Since the statistically based optimal prediction [33] is performed for a truncated KS equation, it provides a convenient setup for comparison with the new and more generally applicable method suggested here.

Another question is how can we apply the RG method to the derivation of a low dimensional approximation for a gKS equation investigated under three scales, i.e., “slow unstable modes”, “slow stable modes”, and “fast stable modes”.

The RG method is based on a natural splitting into linear and nonlinear terms by the variation of constant formula. Recent studies, e.g. by Holden et al. [10, 11], make use of such a splitting via a suitable numerical scheme for equations with Burgers’ nonlinearity. Hence, the reliability and efficiency of the renormalized low dimensional gKS equation motivate the application such numerical splitting strategies to the new reduced equations derived here. Finally, we emphasize that efficient low dimensional approximations are of great interest for numerical scrutiny of long time asymptotes. We shall examine these and related issues in future studies.

**Acknowledgements.** We acknowledge financial support from EPSRC Grant No. EP/H034587, EU-FP7 ITN Multiflow and ERC Advanced Grant No. 247031.

#### REFERENCES

- [1] B. M. Boghosian, C. C. Chow, and T. Hwa. Hydrodynamics of the Kuramoto-Sivashinsky equation in two dimensions. *Phys. Rev. Lett.*, 83(25):5262–5265, 1999.
- [2] L.-Y. Chen, N. Goldenfeld, and Y. Oono. Renormalization group and singular perturbations: Multiple scales, boundary layers, and reductive perturbation theory. *Phys. Rev. E*, 54(1):376–394, 1996.
- [3] A. J. Chorin, A. P. Kast, and R. Kupferman. Optimal prediction of underresolved dynamics. *P. Natl. Acad. Sci. USA*, 95(8):4094–8, 1998.
- [4] C. Chow and T. Hwa. Defect-mediated stability: an effective hydrodynamic theory of spatiotemporal chaos. *Physica D*, 84(3-4):494–512, 1995.
- [5] B. I. Cohen, J. A. Krommes, W. M. Tang, and M. N. Rosenbluth. Nonlinear saturation of the dissipative trapped-ion mode by mode coupling. *Nucl. Fusion*, 16:971–992, 1976.
- [6] C. Duprat, F. Giorgioutti-Dauphiné, D. Tseluiko, S. Saprykin, and S. Kalliadasis. Liquid film coating a fiber as a model system for the formation of bound states in active dispersive-dissipative nonlinear media. *Phys. Rev. Lett.*, 103:234501, 2009.
- [7] C. Foias, M. S. Jolly, I. G. Kevrekidis, and E. S. Titi. On the computation of inertial manifolds. *Phys. Lett. A*, 131(7,8):433–436, 1988.
- [8] C. Foias, B. Nicolaenko, G. R. Sell, and R. Temam. Inertial manifolds for the Kuramoto Sivashinsky equation and an estimate of their lowest dimension. *J. Math. Pure. Appl.*, 67:197–226, 1988.
- [9] L. Giacomelli and F. Otto. New bounds for the Kuramoto-Sivashinsky equation. *Commun. Pur. Appl. Math.*, 58(3):297–318, 2005.
- [10] H. Holden, K. H. Karlsen, N. H. Risebro, and T. Tao. Operator splitting For the KdV equation. *Math. Comput.*, 80(274):821–846, 2011.
- [11] H. Holden, Ch. Lubich, and N. H. Risebro. Operator splitting for partial differential equations with Burgers nonlinearity. *arXiv:1102.4218*, pages 1–12, 2011.
- [12] G. M. Homsy. Model equations for wavy viscous film flow. *Lect. Appl. Math.*, 15:191–194, 1974.
- [13] D. Jackson. *The theory of approximations*. AMS Colloquium Publications, 1930.
- [14] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.
- [15] E. T. Jaynes. Information theory and statistical mechanics II. *Phys. Rev.*, 108:171–190, 1957.
- [16] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1, 1998.
- [17] H. Kalisch and X. Raynaud. On the rate of convergence of a collocation projection of the KdV equation. *ESAIM Math. Model. Numer. Anal.*, 41(1):95–110, 2007.
- [18] S. Kalliadasis, C. Ruyer-Quil, B. Scheid, and M. G. Velarde. *Falling Liquid Films*. Springer Series on Applied Mathematical Sciences, London, 2012.

- [19] Y. Kuramoto and T. Tsuzuki. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Prog. Theor. Phys.*, 55(2):356–369, 1976.
- [20] N. Larkin. Korteweg-de Vries and Kuramoto-Sivashinsky equations in bounded domains. *J. Math. Anal. Appl.*, 297(1):169–185, 2004.
- [21] E. N. Lorenz. Predictability - A problem partly solved. In *ECMWF Seminar on Predictability, 1995*, pages 1–18, Reading, United Kingdom, 1996.
- [22] Y. Maday and A. Quarteroni. Error analysis for spectral approximation of the Korteweg-de Vries equation. *ESAIM Math. Model. Numer. Anal.*, 22(3):499–529, 1988.
- [23] P. Manneville. *Dissipative Structures and Weak Turbulence*. Academic Press, New York, 1990.
- [24] I. Moise and R. Temam. Renormalization group method: Application to Navier-Stokes equation. *Discret. Contin. Dyn. S.*, 6(1):191–210, 2000.
- [25] I. Moise and M. Ziane. Renormalization group method. Applications to partial differential equations. *J. Dyn. Differ. Equ.*, 13(2):275–321, 2001.
- [26] H. Mori. Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.*, 33(3):423–455, 1965.
- [27] F. Otto. Optimal bounds on the Kuramoto-Sivashinsky equation. *J. Funct. Anal.*, 257(7):2188–2245, 2009.
- [28] R. D. Rosenkrantz, editor. *Jaynes, E.T.: Papers on probability, statistics and statistical physics*. Kluwer Academic Publisher, 1989.
- [29] S. Saprykin, E. A. Demekhin, and S. Kalliadasis. Self-organization of two-dimensional waves in an active dispersive-dissipative nonlinear medium. *Phys. Rev. Lett.*, 94:224101, 2005.
- [30] C. E. Shannon. The mathematical theory of communication. *Reprint, The Bell System Technical Journal*, 27(4):379–423, 623–656, 1948.
- [31] B. I. Shraiman. Order, disorder, and phase turbulence. *Phys. Rev. Lett.*, 57(3):325–328, 1986.
- [32] G. Sivashinsky. On self-turbulization of a laminar flame. *Acta Astronaut.*, 6(5-6):569–591, 1979.
- [33] P. Stinis. Stochastic optimal prediction for the Kuramoto-Sivashinsky equation. *Multiscale Model. Simul.*, 2(4):613–638, 2004.
- [34] P. Stinis. A comparative study of two stochastic mode reduction methods. *Physica D*, 213(2):197–213, 2006.
- [35] E. Tadmor. The well-posedness of the Kuramoto-Sivashinsky equation. *SIAM J. Math. Anal.*, 17(4):884, 1986.
- [36] D. Tseluiko, S. Saprykin, C. Duprat, F. Giorgiutti-Dauphiné, and S. Kalliadasis. Pulse dynamics in low-Reynolds-number interfacial hydrodynamics: Experiments and theory. *Physica D*, 239(20-22):2000–2010, 2010.
- [37] D. Tseluiko, S. Saprykin, and S. Kalliadasis. Interaction of solitary pulses in active dispersive-dissipative media. *Proc. Est. Acad. Sci.*, pages 139–144, 2010.
- [38] N. G. van Kampen. Elimination of fast variables. *Phys. Rep.*, 124:69–160, 1985.
- [39] K. G. Wilson. The renormalization group: Critical phenomena and the Kondo problem. *Rev. Mod. Phys.*, 47(4):773–840, 1975.
- [40] R. Zwanzig. Memory effects in irreversible thermodynamics. *Phys. Rev.*, 124(4):983–992, 1961.
- [41] R. Zwanzig. Nonlinear generalized Langevin equations. *J. Stat. Phys.*, 9(3):215–220, 1973.